# Homework 2

Zichen Pan zp2197

Problem 1:

(a) $\hat{\pi} = \arg\max_{\pi} \sum_{i=1}^{n} \ln p(y_i|\pi)$ $\quad p(y_i|\pi) = \pi^{y_i}(1-\pi)^{1-y_i}$

let $l = \sum_{i=1}^{n} \ln p(y_i|\pi) = \sum_{i=1}^{n} \ln\pi \cdot 1_{\{y_i=1\}} + \sum_{i=1}^{n} \ln(1-\pi) \cdot 1_{\{y_i=0\}}$

$= \{\#(y_i=1)\} \cdot \ln\pi + \{\#(y_i=0)\} \cdot \ln(1-\pi)$

$\frac{\partial l}{\partial \pi} = \frac{1}{\pi} \cdot \{\#(y_i=1)\} + \frac{1}{1-\pi}\{\#(y_i=0)\} \equiv 0$, and $\{\#(y_i=0)\} = n - \{\#(y_i=1)\}$

$\Rightarrow \hat{\pi} = \frac{\{\#(y_i=1)\}}{n}$, $i = 1, 2, \cdots, n$

(b) $\hat{\lambda}_{y,d} = \arg\max_{\lambda_{y,d}} \sum_{d=1}^{D} \left( \ln p(\lambda_{y,d}) + \sum_{i=1}^{n} \ln p(x_{i,d}|\lambda_{y_i,d}) \right)$

$= \arg\max_{\lambda_{y,d}} \sum_{d=1}^{D} \left( \ln(\lambda_{y,d} \cdot e^{-\lambda_{y,d}}) + \sum_{i=1}^{n} \ln\left( e^{-\lambda_{y_i,d}} \cdot \frac{\lambda_{y_i,d}^{x_{i,d}}}{x_{i,d}!} \right) \right)$

$= \arg\max_{\lambda_{y,d}} \sum_{d=1}^{D} \left( \ln(\lambda_{y,d}) - \lambda_{y,d} + \sum_{i=1}^{n} \left( -\lambda_{y_i,d} + x_{i,d} \cdot \ln\lambda_{y_i,d} - \ln(x_{i,d}!) \right) \right)$

We leave $y$ arbitrary. Thus we make $y = y_0$.

For a specific $d = d_0$,

$\hat{\lambda}_{y_0,d_0} = \arg\max_{\lambda_{y_0,d_0}} \left( \ln(\lambda_{y_0,d_0}) - \lambda_{y_0,d_0} - \sum_{i=1}^{n} \lambda_{y_i,d_0} \cdot 1_{\{y_i=y_0\}} + \sum_{i=1}^{n} \left( x_{i,d_0} \cdot \ln\lambda_{y_i,d_0} \cdot 1_{\{y_i=y_0\}} \right) \right)$

$\equiv \arg\max_{\lambda_{y_0,d_0}} L$

$\frac{\partial L}{\partial \hat{\lambda}_{y_0,d_0}} = \frac{1}{\lambda_{y_0,d_0}} - 1 - \sum_{i=1}^{n} 1_{\{y_i=y_0\}} + \sum_{i=1}^{n} \frac{x_{i,d_0} \cdot 1_{\{y_i=y_0\}}}{\lambda_{y_0,d_0}} \equiv 0$

$\hat{\lambda}_{y_0,d} = \frac{1 + \sum_{i=1}^{n} (x_{i,d} \cdot 1_{\{y_i=y_0\}})}{1 + \sum_{i=1}^{n} 1_{\{y_i=y_0\}}} = \frac{1 + \sum_{i=1}^{n} (x_{i,d} \cdot 1_{\{y_i=y_0\}})}{1 + \{\#(y_i=y_0)\}}$

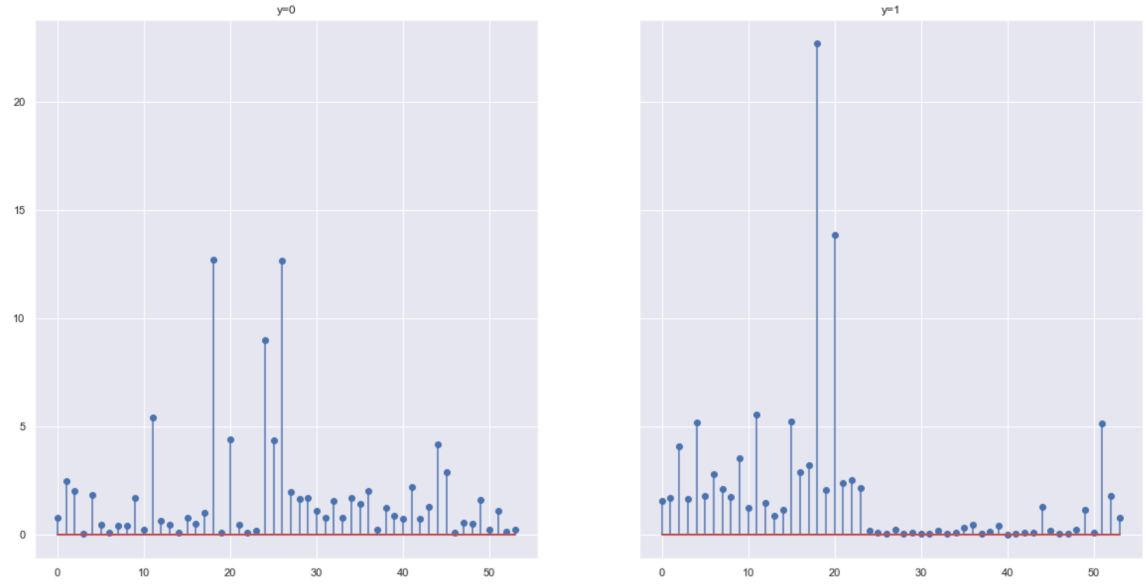
Problem 2

(a)

| Truth / Prediction | 1 | 0 |
|---|---|---|
| 1 | 1703 | 490 |
| 0 | 110 | 2297 |

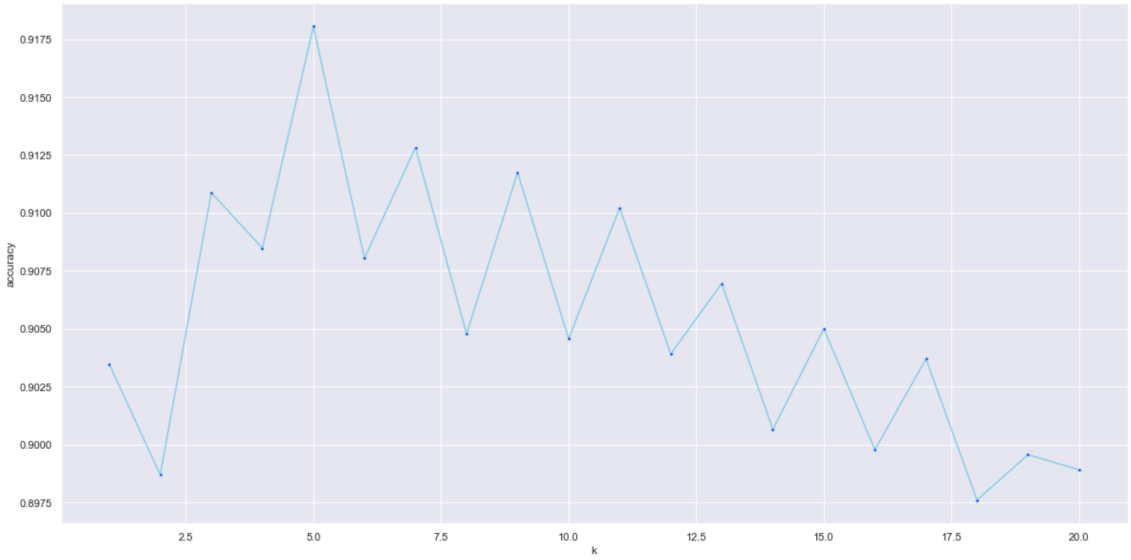
Accuracy = 0.8695652173913043

(b)

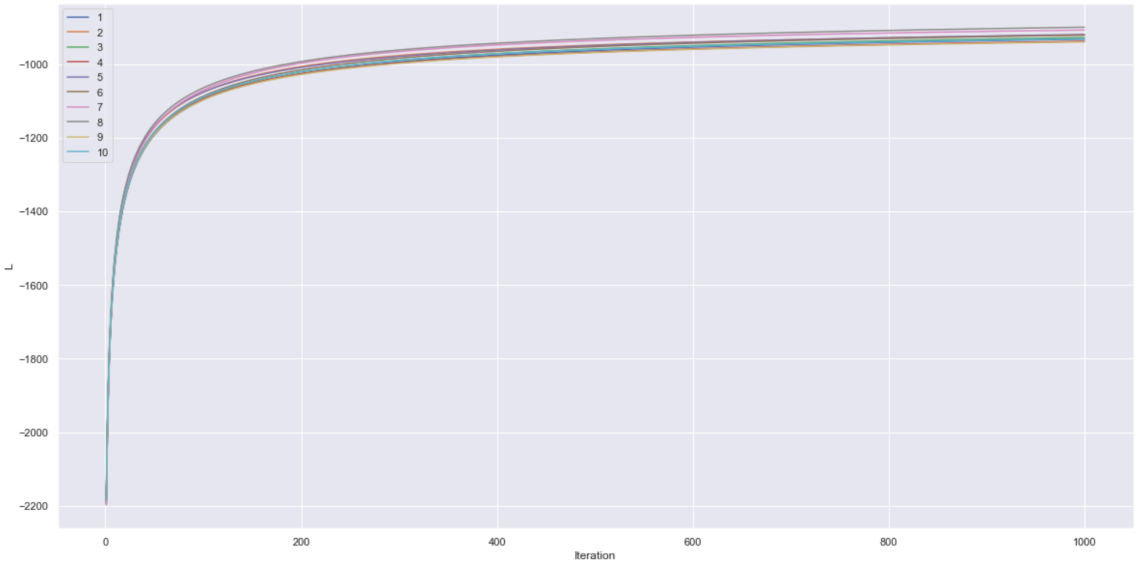



For dimension 16 and 52, the estimated value of lambda for both when y=1 are larger that the counterpart, and y=1 means spam email, which means the word 'free' and character '!' appears more frequently in spam email.

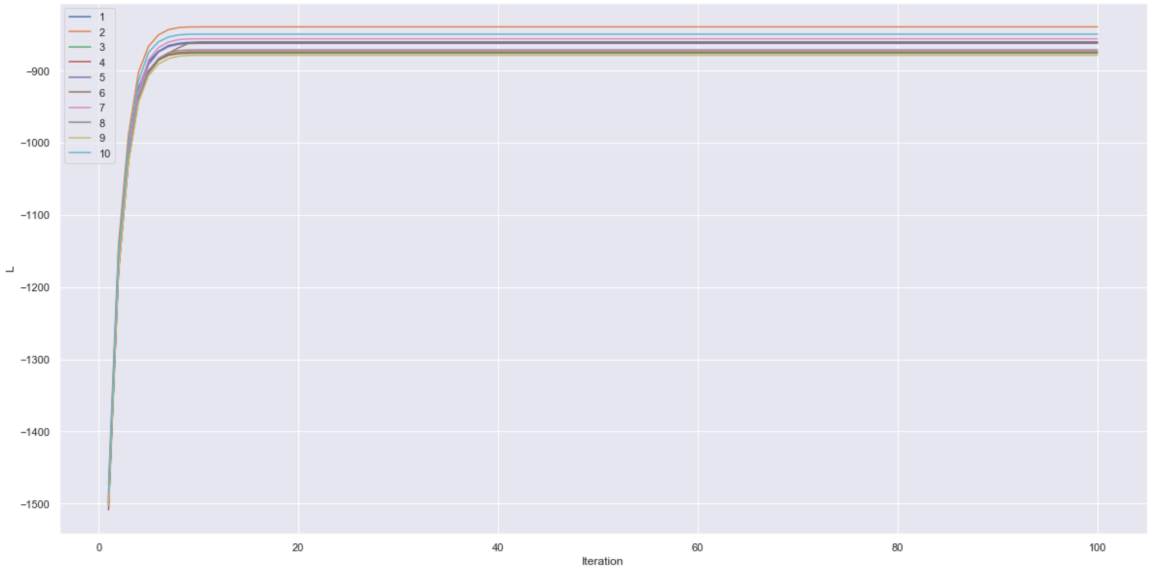
(c)


0.9175
0.9150
0.9125
0.9100
0.9075
0.9050
0.9025
0.9000
0.8975
accuracy
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
k

(d)


1
2
3
4
5
6
7
8
9
10
−1000
−1200
−1400
−1600
−1800
−2000
−2200
L
0
200
400
600
800
1000
Iteration

(e)




(f)

| Truth / Prediction | 1 | 0 |
|---|---|---|
| 1 | 1604 | 144 |
| 0 | 209 | 2643 |

Accuracy = 0.9232608695652174